BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e180280
doi: 10.3897/biss.9.180280

OPEN ACCESS

Conference Abstract

Machine-Actionable Metadata in Practice: Lessons From Automating FAIR Assessment in Plant-Pollinator Datasets

Debora P Drucker[‡], Filipi Miranda Soares[§], Jorrit Poelen[|,¶], José Augusto Salim[#]

‡ Embrapa, Campinas, Brazil
§ INRAE, MISTEA, Montpellier, France
| UC Santa Barbara, Santa Barbara, CA, United States of America
¶ Ronin Institute, Montclair, NJ, United States of America
# Universidade de São Paulo, São Paulo, Brazil

Corresponding author: Debora P Drucker (debora.drucker@embrapa.br)

Received: 24 Nov 2025 | Published: 02 Dec 2025

Citation: Drucker D, Soares F, Poelen J, Salim JA (2025) Machine-Actionable Metadata in Practice: Lessons From Automating FAIR Assessment in Plant-Pollinator Datasets. Biodiversity Information Science and Standards 9: e180280. https://doi.org/10.3897/biss.9.180280

Abstract

Plant-pollinator interactions play a pivotal role in ecosystem functioning and sustainable agriculture. However, plant-polinator datasets are scattered across various networks, in country-specific initiatives, and stored in isolated silos, making them difficult to access by scientists and decision-makers. By promoting the adoption of Findable, Acessible, Interoperable, and Reusable (FAIR) data standards (Wilkinson et al. 2016) across multiple initiatives worldwide, we are working to transform the fragmented nature of these datasets and make data on plant-pollinator interactions widely available.

As the biodiversity community advances towards FAIR data, machine-actionable metadata has emerged as a critical enabler for scalable data assessment, discovery, and reuse. However, while FAIR principles emphasize machine-readability, many datasets are still evaluated manually or lack structured metadata entirely, limiting their integration into global platforms. This study shares practical insights from the WorldFAIR Agricultural Biodiversity Case Study, in which we operationalized machine-actionable FAIR metadata for the review of plant-pollinator interaction datasets (Drucker et al. 2024).

© Drucker D et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

We developed a semi-automated workflow to assist in evaluating datasets against the FAIR principles using tools from the Global Biotic Interactions initiative (GloBI, Poelen et al. 2014). The GloBI bots "Nomer" and "Elton" can read structured metadata from standard vocabularies such as Darwin Core (DwC), Ecological Metadata Language (EML), and the Plant-Pollinator Interactions (PPI) vocabulary. Nomer focuses on taxonomic alignment with several taxonomic catalogues, such as GBIF Backbone and Catalogue of Life. Elton extracts species interactions from datasets of various structures and formats, including DwC-Archives.

By relying on machine-readable metadata, the bots were able to flag inconsistencies, suggest improvements, and generate repeatable reports across pilot projects in Argentina, Brazil, the African continent, Kenya, Colombia, East Africa, Central Asia, and the USA (for example, Elton et al. 2025). This helped researchers assess dataset interoperability without needing full access to the raw data, a crucial feature given legal and institutional access constraints. To make the data review report readable for researchers, GloBI's bots produce a document resembling a data publication, complete with a title, authors, publication date, abstract, introduction, and other relevant sections.

Our results underscore the transformative role of machine-actionable metadata in biodiversity data governance. Automating FAIR assessments not only increases transparency and repeatability but also accelerates the integration of datasets into platforms like GloBI. While human expertise remains essential, tools like Nomer and Elton demonstrate that FAIR assessment can evolve beyond checklists to become dynamic, scalable, and integrated into the data lifecycle. Resources and code are openly available in the online repositories Zenodo*[3] and GitHub*[2], and are summarized in the GloBI platform*[1].

To help alleviate the burden of manually reviewing data as part of scientific publication review, we propose deploying domain-specific, automated data review processes that enable researchers to better understand how to make their data easier to review and reuse. Recognizing that publishing reusable, integrated data remains mostly a manual process, we recommend that plant-pollinator and species interaction datasets be registered with one or more infrastructures (e.g., GloBI, GBIF) to benefit from the domain-specific data review services they offer. We suggest that data publishers continue to collaborate on building, maintaining, and improving similar infrastructures to assess and increase the quality and FAIRness of published scientific data.

Through the adoption of standards such as Ecological Metadata Language, Darwin Core, Plant-Pollinator Interactions Vocabulary, and Relation Ontology, we aim to enhance the understanding of how plant-pollinator interactions contribute to sustaining life on Earth while ensuring that data is easily findable, accessible, interoperable, and reusable for further research and analysis (FAIR).

Keywords

biotic interactions, interoperability, data review, data publication

Presenting author

Debora P Drucker

Presented at

Living Data 2025

Acknowledgements

WorldFAIR Agricultural Biodiversity Case Study collaborators

Grant title

This work was developed as part of the Global cooperation on FAIR data policy and practice (WorldFAIR), funded by the EC HORIZON-WIDERA-2021-ERA-01-41 Coordination and Support Action under Grant Agreement No. 101058393, and we acknowledge the support provided by the project partners and resources.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Drucker DP, Salim JA, Poelen J, et al. (2024) WorldFAIR (D10.3) Agricultural biodiversity FAIR data assessment rubrics. Zenodo https://doi.org/10.5281/zenodo.10719264
- Elton, Nomer, Preston (2025) Versioned Archive and Review of Biotic Interactions and Taxon Names Found within globalbioticinteractions/universityces2023 hash://md5/f5863c1d2b0554419839b3601c580846. Zenodo https://doi.org/10.5281/zenodo.15213046
- Poelen J, Simons J, Mungall C (2014) Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. Ecological Informatics 24 (2014): 148-159. https://doi.org/10.1016/j.ecoinf.2014.08.005
- Wilkinson MD, Dumontier M, Aalbersberg I, et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3: 160018. https://doi.org/10.1038/sdata.2016.18

Endnotes

*1 https://globalbioticinteractions.org/worldfair

*2 https://github.com/globalbioticinteractions

*3 https://zenodo.org/communities/globi-review/